Autofertility and self-compatibility moderately benefit island colonization of plants

Running title: Selfing ability benefits island colonization

Abstract

Aim: The current geographic distribution of species largely reflects colonization success after natural long-distance dispersal or introduction by humans. Plants with selfing ability should have an advantage when establishing on islands where mates and pollinators are limited (Baker's Law). However, high percentages of dioecious and self-incompatible species have been reported for some islands, possibly resulting from post-colonization evolution. Since such evolution is less likely to apply to alien species recently introduced to islands by humans, tests of Baker's Law on islands need to consider both native and naturalized alien species.

Location: Global.

Time period: Undefined.

Major taxa studied: Angiosperms.

Methods: To test whether the colonization of islands is associated with selfing ability (self-compatibility and autofertility), we combined three comprehensive global databases: one on breeding systems of species, one on island and mainland distributions of native species (GIFT) and one on global naturalization of alien plants (GloNAF). We assigned each of a total of 1,752 species, from 161 angiosperm families, as mainland species, island colonists or

island endemics (i.e. species that are restricted to islands). To assess potential relationships between island occurrence and selfing ability of species, we used multinomial logistic regressions.

Results: We found that species with high selfing ability were slightly more likely to be island colonist than mainland species. However, selfing ability did not increase the likelihood of being an island endemic in contrast with mainland species. Among island colonists, selfing ability did not differ between species on oceanic and on continental islands, or between species native to islands and naturalized on islands.

Main conclusions: We performed a comprehensive test of Baker's Law by considering many angiosperm families, using continuous metrics of self-compatibility and autofertility, and including both native and naturalized species. We provide global evidence that high selfing ability may foster island colonization of angiosperms.

Keywords: Breeding system; exotic species; invasions; island biogeography; mating system; reproductive biology.

Introduction

The geographic distribution of species on oceanic islands reflects colonization success after long-distance dispersal, subsequent evolution and radiation, and – more recently – introduction by humans (MacArthur & Wilson, 1967; Whittaker & Fernández-Palacios, 2007; Moser *et al.*, 2018). Although identifying the determinants of successful colonization is challenging, especially after rare events of long-distance dispersal, geographic patterns in species traits can provide valuable insights into determinants of colonization success. Baker's Law refers to the advantage of species with a capacity for uniparental reproduction when colonizing new habitat after long-distance dispersal and when pollinators are limited (Baker, 1955, 1967). Baker's Law may apply to metapopulation dynamics, natural range expansions, biological invasions and particularly the colonization of islands (Pannell *et al.*, 2015). When plant propagules arrive on an island, the resulting plants are isolated from their source populations by oceans, and mates are likely to be limiting. Under such circumstance, self-compatible species should have an advantage because they can reproduce from a single individual. Moreover, as animal species in general, and major groups of pollinators in particular, are less frequent on islands than in mainland regions (Barrett, 1996), autofertile plants (i.e. self-compatible plants being able to self-pollinate) should have a particularly strong advantage when establishing on islands.

Baker's Law has often been challenged due to contradictory findings. Many studies found a prevalence of self-compatible species on islands, supporting Baker's Law (Bernardello *et al.*, 2001; Chamorro *et al.*, 2012; Lord, 2015), while other studies found relatively high frequencies of self-incompatible and dioecious species in the floras of oceanic islands (Carlquist, 1966; Bawa, 1982; Sakai *et al.*, 1995). This apparent contradiction might result from the evolution of species reproductive traits following establishment (Pannell *et al.*,

2015). On the one hand, the paucity of insect pollinators on islands has been suggested to favor the evolution of floral traits, such as reductions in flower size and anther-stigma separation, that increase selfing in self-compatible plants (Barrett, 1996). On the other hand, outcrossing should be advantageous for establishing in new environments given that it avoids inbreeding depression and widens the gene pool, thus allowing for local adaptation (Charlesworth & Charlesworth, 1987). The selection of reduced selfing ability and enhanced outcrossing mechanisms such as dioecy, herkogamy and dichogamy may therefore be important for the subsequent radiation and diversification of species following establishment on islands (Barrett, 1996).

Recently, Grossenbacher *et al.* (2017) found a greater proportion of self-compatible species on islands than in mainland regions in a study on >1,500 species. This study provides strong support for Baker's Law, but it was restricted to three angiosperm families, and did not look at autofertility, i.e. the ability of self-compatible plants to self-pollinate. Moreover, it only differentiated between mainland and island species, both including and excluding the island endemics, and did not test whether self-compatibility differed between island endemics and non-endemics (i.e. species that are not restricted to islands). In addition to differentiating between endemics and non-endemics among island species, testing Baker's Law would benefit from including naturalized alien species, which established on islands relatively recently (i.e. in the last few centuries), since their potential for major evolutionary post-arrival transitions is minimal. More than 13,000 alien plant species have been introduced by humans to new regions where they have become part of the local flora (van Kleunen *et al.*, 2015; Pyšek *et al.*, 2017), and most of these species naturalization events happened in the last two centuries (Seebens *et al.*, 2017; Seebens *et al.*, 2018). Therefore, the likelihood for evolutionary changes in self-compatibility and autofertility of these naturalized species is small. However, no study testing the applicability of Baker's Law in the colonization of

islands has included naturalized species so far. We hence lack a comprehensive overview of the role of both self-compatibility and autofertility of species across a representative set of angiosperm families of angiosperms in island colonization.

The strength of the association between island colonization and selfing ability may depend on island geological history. As continental islands have been connected to the mainland in the past, mates and pollinators might be less likely to be lacking than on the usually much younger and more remote oceanic islands, which have never been connected to the mainland and where all species were new colonizers or their descendants. In line with this, Grossenbacher *et al.* (2017) found that for Asteraceae, Brassicaceae and Solanaceae, self-compatible species occurred more often on islands than self-incompatible ones did, and that this effect was stronger for oceanic than for continental islands. Whether this also holds true across other angiosperm families remains to be tested.

Selfing ability and other species characteristics associated with breeding-system and colonization success are continuous (Raduski *et al.*, 2012; Razanajatovo *et al.*, 2016), and treating them as such should provide a rigorous test of the relationship between breeding-system and island colonization of plant species. Quantitative metrics of breeding-system traits should be more informative than the commonly used qualitative measures, such as self-compatible vs. self-incompatible, when addressing general ecological questions related to plant reproductive strategies. However, previous studies testing the role of breeding systems in island colonization mainly relied on qualitative measures. Furthermore, other plant characteristics such as life history can be associated with the biogeographic patterns in plant breeding systems (Razanajatovo *et al.*, 2016; Moeller *et al.*, 2017). Therefore, in addition to using continuous metrics of self-compatibility and autofertility, it is important to account for life history when testing the relationship between island colonization and selfing ability.

Here we combined three comprehensive global databases to test whether plant species with high self-compatibility and autofertility were more likely to occur on islands, as predicted by Baker's Law. We assumed that species that colonized islands, either naturally or after introduction by humans, arrived from the mainland. Thus, we considered that species currently occurring on the mainland and absent from island floras have failed to colonize islands (hereafter referred to as mainland species). Species that occur both on the mainland and on islands are likely to have succeeded in colonizing an island from the mainland (island colonists). Finally, species occurring only on islands are likely to have evolved there after arrival of a colonist ancestor (island endemics). We explicitly consider these different types of plant distributions (i.e. mainland species, island colonists, island endemics), and ask: 1) Are plant species with high selfing ability more likely to occur on both the mainland and islands than only on the mainland (i.e. to be island colonists)? 2) Are plant species with high selfing ability more likely to occur only on islands than only on the mainland (i.e. to be island endemics)? 3) Among the island colonists, is there a difference in selfing ability between species on oceanic islands and those on continental islands? 4) Is there a difference in selfing ability between species native to islands and those naturalized on islands?

Methods

Species data

To test whether the colonization of islands by angiosperm species, irrespective of whether they are native or naturalized, is associated with selfing ability, we combined three comprehensive global plant databases: one on breeding systems of species, one on island and mainland distributions of native species (GIFT; gift.uni-goettingen.de) and one on global naturalization success of alien species (GloNAF; van Kleunen *et al.*, 2015). To obtain

131 information on the selfing ability of each species in this study, we used a global database on
132 breeding systems of angiosperms (for details on the compilation see Razanajatovo *et al.*,
133 2016). Briefly, selfing-ability indices were calculated using fruit set and seed production after
134 different breeding-system treatments. First, a self-compatibility index was obtained by
135 dividing the outcome of self-pollination by the outcome of outcross-pollination treatments.
136 Second, an autofertility index was obtained by dividing the outcome of pollinator exclusion
137 by that of outcross-pollination treatments. This database includes self-compatibility and
138 autofertility indices calculated for 1,752 angiosperm species from 161 families from all
139 continents except Antarctica.

140 To obtain information on the native distribution of each species with quantitative
141 breeding-system data on islands and mainlands, we used the Global Inventory of Floras and
142 Traits (GIFT; (Weigelt *et al.*, 2017); http://gift.uni-goettingen.de), which to date included
143 regional plant species lists for 1,636 islands and 993 mainland regions globally. Depending on
144 the plant taxonomic group, GIFT covers between 70% and 100% of the terrestrial surface of
145 the world for native plant species. Because GIFT does not have a global coverage for all
146 species, 61.5 % of the species in this study did not have complete global island/mainland
147 distribution information. For these species, if we had only information on their occurrence on
148 islands, we could not rule out the possibility that they also occur in mainland regions, and if
149 we had only information on their occurrence in mainland regions, we could not rule out the
150 possibility that they occur also on islands. Nevertheless, given the comprehensive coverage of
151 islands in GIFT, it is unlikely that a species occurs on any island, especially oceanic ones, if it
152 is not listed in the floras of the islands included in GIFT. For species occurring on islands, we
153 also obtained information on whether they occur on oceanic and/or continental islands from
154 GIFT. In an additional analysis using a more conservative approach, we excluded species

known to occur in either island or mainland regions but for which information was missing about their occurrence in the other region type.

To obtain information on the naturalized distribution of the species included in this study, we used the Global Naturalized Alien Flora (GloNAF version 1.1; van Kleunen *et al.*, 2015; Pyšek *et al.*, 2017). This database includes 13,168 naturalized plant species and covers 362 islands and 481 mainland regions globally. We checked for each naturalized species (i.e. listed in GloNAF; n=498), whether it is naturalized on islands and/or in mainland regions.

Island and mainland occurrence of native and naturalized plant species

Using the information on the geographic distribution of the native and naturalized species from the GIFT and the GloNAF databases, we assigned each species with breeding-system data to one of three categories according to their occurrence in mainland and island regions (mainland species, island colonists, island endemics). When a species was known to occur on an island but information was missing for its occurrence in mainland regions, we considered the species as not occurring in mainland regions. When a species was known to occur in mainland regions but information about island occurrences was missing, we assumed that they do not occur on islands.

In a first categorization, we considered each island regardless of whether it is an oceanic island or not. If the species was known to occur only in mainland regions as native or naturalized, we assigned it to mainland species. If a species was known to occur in mainland regions and at least in one island region, we assigned it to island colonists. If a species was known to occur in at least one island region and not known to occur in any mainland region, we assigned it to island endemics. In a second categorization, as continental islands may have been colonized before they became islands, we assigned species distributions according to their occurrence on oceanic islands only.

179 **Statistical analysis**

180 To quantitatively assess selfing ability of species, we used four selfing-ability indices:
181 two self-compatibility indices and two autofertility indices calculated using fruit set and seed
182 production (see (Razanajatovo *et al.*, 2016) for details on the calculation of the different
183 indices). To assess potential relationships between island occurrence and selfing ability of
184 species, we used multinomial logistic regressions (Ntzoufras, 2011). We used species
185 distribution on the mainland and islands as a multi-categorical response variable. The
186 response variable had the following three categories: (1) mainland species (species occurring
187 only on the mainland), (2) island colonists (species occurring on both the mainland and
188 islands), and (3) island endemics (species occurring only on islands). As explanatory
189 variables, we used the self-compatibility or the autofertility index of species, scaled to a mean
190 of zero and a standard deviation of one, their life history (annual/biennial or perennial,
191 corresponding to monocarpic and polycarpic, respectively), and the interaction between self-
192 compatibility or autofertility and life history. To account for non-independence of species due
193 to evolutionary relatedness, we used family as a random factor. We ran multinomial logistic
194 regressions in WinBUGS (Lunn *et al.*, 2000) from R version 2.15.3 (R Core Team, 2012),
195 using the 'R2WinBUGS' package (Sturtz *et al.*, 2005). The model parameters were estimated
196 from Gibbs sampling of Markov chain Monte Carlo (MCMC) using three chains. For each
197 dataset, 50,000 iterations were run after an initial burn-in phase of 1,000 iterations. To reduce
198 the correlation between consecutive samples, only every fifth value was taken per chain. To
199 assess the goodness-of-fit of the models, we compared the model predictions with the data
200 graphically (Korner-Nievergelt *et al.*, 2015). Separately for each category of the response
201 variable (mainland species, island colonists and island endemics), we plotted for each species
202 the probability to be a certain category (i.e. the model predictions) on the x-axis, and the
203 observations (0=not this category, 1=this category) on the y-axis. We added class-wise means

of the observations within classes of width 0.1 of the model predictions. We also added the y=x line. If the model fits the data well, the data would be on average equal to the model predictions, i.e. the class-wise means should be close to the y=x line.

To test whether autofertility is more strongly related to island occurrence than self-compatibility is, we used a subset of species for which both self-compatibility and autofertility indices were available. We ran similar multinomial logistic regressions as described above, and compared the Deviance Information Criterion (DIC) of the model with self-compatibility to that with autofertility as explanatory variable. Additionally, we ran multinomial logistic regressions in which we included both self-compatibility and autofertility indices of species, scaled to a mean of zero and a standard deviation of one, the interaction between self-compatibility and autofertility, species' life history (annual/biennial or perennial), the interaction between self-compatibility and life history, and the interaction between autofertility and life history as explanatory variables.

To test among the island colonists whether there is a difference in selfing ability between species that occur on oceanic islands and those that occur on continental islands, we fitted logistic regressions using the glmer function of the 'lme4' package (Bates *et al.*, 2015) in R. As the binary response variable, we used whether a species occurs on oceanic islands or not (yes/no). As explanatory variables, we included the self-compatibility or the autofertility index of the species, scaled to a mean of zero and a standard deviation of one, their life history (annual/biennial or perennial), and the interaction between self-compatibility or autofertility and life history. To test whether there is a difference in selfing ability between species native to islands and species naturalized on islands, we built logistic regressions, using the same terms as in the previous model. As the binary response variable, we used whether a species is naturalized on islands or not (yes/no). In these models, we also used family as a

random factor. We tested for significance of the fixed terms using log-likelihood ratio tests (Zuur *et al.*, 2009).

Results

Of the 1,752 plant species included in this study, 1,129 are mainland species, 566 are island colonists, and 57 are island endemics. Of the 623 species occurring on islands, 242 occur on at least one oceanic island, 463 are native on islands, and 285 are naturalized on islands. Because a species can be native to an island and naturalized on another island, the total number of species native to islands and that of species naturalized on islands do not add up to the total number of species occurring on islands. Since annual and biennial species were more likely to be island colonists, and less likely to be mainland species (Table 1; Figure S1 in Supporting Information), the results on self-compatibility and autofertility below have been corrected for differences in life history. Graphical assessment of model fits showed that the models fit the data well (Figures S2-S6).

Association of island colonists and island endemics with selfing ability

When selfing-ability indices based on fruit set were used, we found that species with high self-compatibility and autofertility indices were slightly more likely to be island colonist than mainland species (Table 1). An increase of 0.405 and 0.353 (one standard deviation) in self-compatibility and autofertility index of a species increased its odds of being an island colonist by a factor of 1.21 and 1.43, respectively (Figures 1a, 2a, and 3). Self-compatibility and autofertility, however, did not increase the likelihood of being an island endemic (Table 1; Figures 1b, 2b and 3). When selfing-ability indices based on seed production (n=419 for self-compatibility and n=284 for autofertility index) instead of fruit set (n=1184 for self-compatibility and n=868 for autofertility index) were used, the results were slightly different

(compare Figures 1a-b and 1c-d; 2a-b and 2c-d). Self-compatibility still increased the likelihood of being an island colonist (Figures 1c and 3), but there was no relationship between the likelihood of being an island colonist and autofertility (Table 1; Figures 2c and 3). Overall, there was no interaction between self-compatibility or autofertility and life history (Table 1).

Association of the occurrence of species on oceanic vs. continental islands with selfing ability

When only oceanic islands were considered, there was a weak but detectable positive relationship between the likelihood of being an island colonist (instead of a mainland species) and autofertility, but self-compatibility did not increase the likelihood of being an island colonist (compare Tables 1 and S1; Figures 1-2 and S7). The positive relationship between the likelihood of being an island colonist and autofertility tended to be weaker for annual and biennial species than for perennial species, as indicated by an interaction between autofertility index based on seed production and life history (Table S1). Among the island colonists, self-compatibility and autofertility did not differ between species that occur on oceanic islands and species that occur on continental islands (Table 2).

Association of the occurrence of native vs. naturalized species on islands with selfing ability

Among the island colonists, species native to islands and species naturalized on islands did not differ in their degrees of self-compatibility and autofertility (Table 3).

Autofertility vs. self-compatibility as a driver of island occurrence

Within the subset of species for which both self-compatibility and autofertility indices were available, when indices were based on fruit set, we found that there was a tendency for species with high self-compatibility and autofertility indices to be island colonists rather than mainland species. However, the Deviance Information Criterion (DIC) of the model with autofertility was lower than that of the model with self-compatibility (Table S2), indicating a

better fit. When indices based on seed production were used, there was no relationship between the likelihood of being an island colonist and autofertility (Table S2). When both self-compatibility and autofertility indices were included in a model, there was still a positive relationship between the likelihood of being an island colonist and autofertility, at least so in the analysis with indices based on fruit set (Table S3).

Species with incomplete data on mainland and island distribution excluded

In an additional analysis using a more conservative approach, when species with incomplete data on mainland and island distribution were excluded, the results were slightly different. Then, species with high autofertility were slightly more likely to be island colonist than mainland species, but there was no relationship between the likelihood of being an island colonist and self-compatibility (compare Tables 1 and S4; Figures 1-2 and S8). Moreover, the positive relationship between the likelihood of being an island colonist and autofertility tended to be stronger for annual and biennial species than for perennial species, as indicated by an interaction between autofertility index based on fruit set and life history (Table S4).

Discussion

This is the first test of Baker's Law considering a large set of angiosperm families, using continuous metrics of self-compatibility and autofertility, and including both native and naturalized species. Despite some variation in the results of the different tests and subsets of data, plant species with high selfing ability, regardless of being native or naturalized, tended generally to be island colonists rather than mainland species, supporting Baker's Law. We found similar patterns of association between island occurrence and selfing ability, whether we considered all islands or only oceanic islands.

We found that both self-compatible and autofertile plant species were more likely to be island colonists. Even if self-compatibility and autofertility are not independent, the model with autofertility gave a better fit than the one with self-compatibility (Table S2), and in the model with both self-compatibility and autofertility, there was still a positive relationship between the likelihood of being an island colonist and autofertility (Table S3). This suggests that autofertility, which requires both autonomous self-fertilization and self-compatibility (or apomixis), is a stronger driver of island colonization than self-compatibility alone. These findings corroborate previous ones on the role of autofertility in the colonization of new regions in different contexts. For example, in North America, the over-representation of autonomously selfing plants in populations of *Campanula americana* at the northern and the western range edges has been suggested to be the result of post-glacial recolonization from southern refugia (Koski *et al.*, 2017). Although a previous review by Pannell *et al.* (2015) proposed a restriction of the scope of Baker's Law to the consequences of mate rather than pollinator limitation, besides the benefit of overcoming mate limitation by self-compatibility, autofertile species also have the advantage of overcoming pollinator limitation when establishing in new regions, especially on islands where the major groups of pollinators are usually scarce.

We did not find an association between selfing ability and island endemics in contrast with mainland species. With the exception of a few relict endemics (i.e. species that went extinct elsewhere), island endemic species might have arisen from different modes of speciation such as cladogenetic and anagenetic speciation (Stuessy *et al.*, 1990). Whichever is the mode of speciation, island endemics might have evolved different breeding systems compared to their mainland ancestors. The colonizers that the endemics evolved from may also have gone through mate and pollinator limitation filters, and, similarly to the island colonists, have had a higher selfing ability than mainland species. A subsequent evolution of

reduced selfing ability and more outcrossing or dioecy compared to the mainland species might then have balanced out the influence of the filtering. Nevertheless, as island endemics accounted only for a relatively small number of species in our database (n=57), the question whether island endemic species have generally evolved higher or reduced selfing ability remains to be tested more rigorously.

We found that selfing ability did not differ between native and naturalized island colonists suggesting that high selfing ability might also help alien species to become naturalized on islands where suitable mates and pollinators are more limiting than on the mainland. Some plants that were introduced by humans to both island and mainland regions from mainland regions have previously been shown to exhibit higher selfing ability on islands than on the mainland, as for *Nicotiana glauca* on two of the California Channel Islands compared to on the California mainland (Schueller, 2004). *Nicotiana glauca* is pollinated by hummingbirds and sunbirds in regions where it was introduced and where such birds are present, but it is largely selfing in regions where bird pollinators are absent, such as on the island of Tenerife (Ollerton *et al.*, 2012). On the other hand, multiple introductions of alien species have been shown to have alleviated mate limitation during establishment (Dlugosch & Parker, 2008; Pannell *et al.*, 2015). In line with these case studies and the previous finding that species with selfing ability were more likely to become naturalized (Razanajatovo *et al.*, 2016), our results also suggest that selfing ability may benefit the establishment of alien species on islands.

Correlations among and trade-offs between different species traits complicate the detectability of Baker's Law. Life history and growth form have been shown to play a major role in the geographic distribution of plant species (Razanajatovo *et al.*, 2016; Moeller *et al.*, 2017). We found that annual and biennial species were more likely to be island colonists. Perennial species have more time to reproduce than annual and biennial species, and might

347 therefore more easily overcome mate and pollinator limitations (Pannell *et al.*, 2015).
348 Nevertheless, after accounting for life history, we still found an association between selfing
349 ability and island occurrence. Other reproductive traits, associated or not with selfing ability,
350 can further contribute to variation in species distribution patterns. For example, as polyploidy
351 can be important for the evolution and diversification of the colonists' progeny, polyploid
352 species may have larger ranges than the diploid ones (Lowry & Lester, 2006). Clonality can
353 also help species without selfing ability to establish populations in new environments (Pyšek,
354 1997; Vallejo-Marín & O'Brien, 2007), and this trait is also implied in the capacity for
355 uniparental reproduction proposed by Baker's Law (Pannell *et al.*, 2015). Furthermore, traits
356 related to dispersal ability can confer advantages in the colonization of islands. The relatively
357 high frequencies of dioecious species on islands may, for example, be linked to fleshy and
358 many-seeded fruits (Vamosi *et al.*, 2007), increasing the number of seeds dispersed by seed
359 dispersal mutualists. Therefore, both dispersal and breeding system traits can affect
360 colonization success.

361 Our results were partly different depending on whether we included or excluded the
362 species with incomplete data on island/mainland distribution (compare Tables 1 and S4). If
363 excluded, the association between occurrence on islands and selfing ability was weaker (i.e.
364 less often positive). However, this more conservative dataset may be biased towards island
365 species, as the GIFT database started with the compilation of island floras (Weigelt *et al.*,
366 2017). Because only a few mainland species remained in this dataset, i.e. most species were
367 island colonists, the comparison between island and mainland occurrences might have been
368 less powerful than when all species were included. Nevertheless, whenever we found an
369 association between island occurrence and selfing ability, it was positive, and thus in line with
370 Baker's Law (Table S4).

Our results were also partly different depending on whether we used selfing ability indices calculated by using fruit set or seed production. Although seed production would be a preferable proxy for maternal fitness, it is practically challenging to measure seed production especially for non-herbaceous species. For this reason, compared to seed production, fruit set has been the most commonly measured variable for reproductive success in pollination studies (Knight et al., 2005; Razanajatovo et al., 2016). Nevertheless, fruit set and seed production can show similar patterns, as was found for pollen limitation (Knight et al., 2005). For the subset of species for which we had selfing ability indices based on fruit set and seed production, the two indices were highly correlated (self-compatibility index: Pearson's $r=0.833$, $p<0.0001$, $n=384$, Figure S9a; autofertility index: Pearson's $r=0.835$, $p<0.0001$, $n=263$, Figure S9b). Because of the much larger sample sizes in our analyses with selfing ability indices based on fruit set compared to those based on seed production (e.g. n=1184 vs. 419 for self-compatibility index, and n=868 vs. 284 for autofertility index in the non-conservative dataset), the analyses with indices based on fruit set should have more power (van Kleunen et al., 2014). Furthermore, when we found an association between island occurrence and selfing ability indices based on seed production, it was positive (Figures 1, 2 and S7), thus supporting Baker's Law.

Conclusion

Our comprehensive test of Baker's Law shows that selfing ability may influence the colonization of islands by plants, both naturally after long-distance dispersal and with the help of humans. The ability to reproduce is critical for establishment in new environments, and reproductive traits involving the breeding system can be vital. The role of breeding system might be especially notable in the colonization of islands, as colonists have fewer opportunities to find mates than on the mainland, and on islands the major groups of

pollinators are often impoverished. Our results suggest that both mate- and pollinator limitations may act as filters to the successful colonization of islands by plant species.

References

Baker, H.G. (1955) Self-compatibility and establishment after long distance dispersal. *Evolution*, 9, 347-349.

Baker, H.G. (1967) Support for Baker's law as a rule. *Evolution*, 21, 853-856.

Barrett, S.C. (1996) The reproductive biology and genetics of island plants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 351, 725-733.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1-48.

Bawa, K. (1982) Outcrossing and the incidence of dioecism in island floras. *The American Naturalist*, 119, 866-871.

Bernardello, G., Anderson, G.J., Stuessy, T.F. & Crawford, D.J. (2001) A survey of floral traits, breeding systems, floral visitors, and pollination systems of the angiosperms of the Juan Fernandez Islands (Chile). *The Botanical Review*, 67, 255-308.

Carlquist, S. (1966) The biota of long-distance dispersal. IV. Genetic systems in the floras of oceanic islands. *Evolution*, 20, 433-455.

Chamorro, S., Heleno, R., Olesen, J.M., McMullen, C.K. & Traveset, A. (2012) Pollination patterns and plant breeding systems in the Galápagos: a review. *Annals of Botany*, 110, 1489-1501.

Charlesworth, D. & Charlesworth, B. (1987) Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics*, 18, 237-268.

417 Dlugosch, K. & Parker, I. (2008) Founding events in species invasions: genetic variation,
418 adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, 17,
419 431-449.
420 Grossenbacher, D.L., Brandvain, Y., Auld, J.R., Burd, M., Cheptou, P.-O., Conner, J.K.,
421 Grant, A.G., Hovick, S.M., Pannell, J.R., Pauw, A., Petanidou, T., Randle, A.M.,
422 Rubio de Casas, R., Vamosi, J., Winn, A., Igic, B., Busch, J.W., Kalisz, S. &
423 Goldberg, E.E. (2017) Self-compatibility is over-represented on islands. *New*
424 *Phytologist*, 215, 469-478.
425 Knight, T.M., Steets, J.A., Vamosi, J.C., Mazer, S.J., Burd, M., Campbell, D.R., Dudash,
426 M.R., Johnston, M.O., Mitchell, R.J. & Ashman, T.-L. (2005) Pollen limitation of
427 plant reproduction: pattern and process. *Annual Review of Ecology, Evolution, and*
428 *Systematics*, 36, 467-497.
429 Korner-Nievergelt, F., Roth, T., Von Felten, S., Guélat, J., Almasi, B. & Korner-Nievergelt,
430 P. (2015) *Bayesian data analysis in ecology using linear models with R, BUGS, and*
431 *Stan*. Academic Press, Amsterdam.
432 Koski, M.H., Grossenbacher, D.L., Busch, J.W. & Galloway, L.F. (2017) A geographic cline
433 in the ability to self-fertilize is unrelated to the pollination environment. *Ecology*, 98,
434 2930-2939.
435 Lord, J.M. (2015) Patterns in floral traits and plant breeding systems on Southern Ocean
436 Islands. *AoB Plants*, 7, plv095.
437 Lowry, E. & Lester, S. (2006) The biogeography of plant reproduction: potential determinants
438 of species' range sizes. *Journal of Biogeography*, 33, 1975-1982.
439 Lunn, D.J., Thomas, A., Best, N. & Spiegelhalter, D. (2000) WinBUGS-a Bayesian modelling
440 framework: concepts, structure, and extensibility. *Statistics and Computing*, 10, 325-
441 337.

MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, New Jersey.

Moeller, D.A., Briscoe Runquist, R.D., Moe, A.M., Geber, M.A., Goodwillie, C., Cheptou, P.-O., Eckert, C.G., Elle, E., Johnston, M.O., Kalisz, S., Ree, R.H., Sargent, R.D., Vallejo-Marin, M. & Winn, A.A. (2017) Global biogeography of mating system variation in seed plants. *Ecology Letters*, 20, 375-384.

Moser, D., Lenzner, B., Weigelt, P., Dawson, W., Kreft, H., Pergl, J., Pyšek, P., van Kleunen, M., Winter, M., Capinha, C., Cassey, P., Dullinger, S., Economo, E.P., García-Díaz, P., Guénard, B., Hofhansl, F., Mang, T., Seebens, H. & Essl, F. (2018) Remoteness promotes biological invasions on islands worldwide. *Proceedings of the National Academy of Sciences*, DOI: 10.1073/pnas.1804179115.

Ntzoufras, I. (2011) *Bayesian modeling using WinBUGS*. John Wiley & Sons, New Jersey.

Ollerton, J., Watts, S., Connerty, S., Lock, J., Parker, L., Wilson, I., Schueller, S.K., Nattero, J., Cocucci, A.A. & Izhaki, I. (2012) Pollination ecology of the invasive tree tobacco Nicotiana glauca: comparisons across native and non-native ranges. *Journal of Pollination Ecology*, 9, 85-95.

Pannell, J.R., Auld, J.R., Brandvain, Y., Burd, M., Busch, J.W., Cheptou, P.O., Conner, J.K., Goldberg, E.E., Grant, A.G. & Grossenbacher, D.L. (2015) The scope of Baker's law. *New Phytologist*, 208, 656-667.

Pyšek, P. (1997) Clonality and plant invasions: can a trait make a difference. *The ecology and evolution of clonal plants*, eds. de Kroon H. & van Groenendael J., Backhuys Publishers Leiden, pp. 405-427.

Pyšek, P., Pergl, J., Essl, F., Lenzner, B., Dawson, W., Kreft, H., Weigelt, P., Maurel, N. & van Kleunen, M. (2017) Naturalized alien flora of the world: species diversity, taxonomic and phylogenetic patterns, geographic distribution and global hotspots of plant invasion. *Preslia*, 89, 203-274.

R Core Team (2012) *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raduski, A.R., Haney, E.B. & Igić, B. (2012) The expression of self-incompatibility in angiosperms is bimodal. *Evolution*, 66, 1275-1283.

Razanajatovo, M., Maurel, N., Dawson, W., Essl, F., Kreft, H., Pergl, J., Pyšek, P., Weigelt, P., Winter, M. & van Kleunen, M. (2016) Plants capable of selfing are more likely to become naturalized. *Nature Communications*, 7, 13313.

Sakai, A.K., Wagner, W.L., Ferguson, D.M. & Herbst, D.R. (1995) Origins of dioecy in the Hawaiian flora. *Ecology*, 76, 2517-2529.

Schueller, S.K. (2004) Self-pollination in island and mainland populations of the introduced hummingbird-pollinated plant, *Nicotiana glauca* (Solanaceae). *American Journal of Botany*, 91, 672-681.

Seebens, H., Blackburn, T.M., Dyer, E.E., Genovesi, P., Hulme, P.E., Jeschke, J.M., Pagad, S., Pyšek, P., Winter, M., Arianoutsou, M., Bacher, S., Blasius, B., Brundu, G., Capinha, C., Celesti-Grapow, L., Dawson, W., Dullinger, S., Fuentes, N., Jäger, H., Kartesz, J., Kenis, M., Kreft, H., Kühn, I., Lenzner, B., Liebhold, A., Mosena, A., Moser, D., Nishino, M., Pearman, D., Pergl, J., Rabitsch, W., Rojas-Sandoval, J., Roques, A., Rorke, S., Rossinelli, S., Roy, H.E., Scalera, R., Schindler, S., Štajerová, K., Tokarska-Guzik, B., van Kleunen, M., Walker, K., Weigelt, P., Yamanaka, T. & Essl, F. (2017) No saturation in the accumulation of alien species worldwide. *Nature Communications*, 8, 14435.

Seebens, H., Blackburn, T.M., Dyer, E.E., Genovesi, P., Hulme, P.E., Jeschke, J.M., Pagad, S., Pyšek, P., van Kleunen, M., Winter, M., Ansong, M., Arianoutsou, M., Bacher, S., Blasius, B., Brockerhoff, E., Brundu, G., Capinha, C., Causton, C., Celesti-Grapow, L., Dawson, W., Dullinger, S., Economo, E., Fuentes, N., Guénard, B., Jäger, H., Kartesz, J., Kenis, M., Kühn, I., Lenzner, B., Liebhold, A., Mosena, A., Moser, D.,

494 Nentwig, W., Nishino, M., Pearman, D., Pergl, J., Rabitsch, W., Rojas-Sandoval, J.,
495 Roques, A., Rorke, S., Rossinelli, S., Roy, H.E., Scalera, R., Schindler, S., Štajerová,
496 K., Tokarska-Guzik, B., Walker, K., Ward, D.F., Yamanaka, T. & Essl, F. (2018)
497 Global rise in emerging alien species results from increased accessibility of new
498 source pools. *Proceedings of the National Academy of Sciences*, 201719429.
499 Stuessy, T.F., Crawford, D.J. & Marticorena, C. (1990) Patterns of phylogeny in the endemic
500 vascular flora of the Juan Fernandez Islands, Chile. *Systematic Botany*, 15, 338-346.
501 Sturtz, S., Ligges, U. & Gelman, A. (2005) R2WinBUGS: A Package for Running WinBUGS
502 from R. *Journal of Statistical Software*, 12, 1-16.
503 Vallejo-Marín, M. & O'Brien, H.E. (2007) Correlated evolution of self-incompatibility and
504 clonal reproduction in Solanum (Solanaceae). *New Phytologist*, 173, 415-421.
505 Vamosi, J.C., Zhang, Y. & Wilson, W.G. (2007) Animal dispersal dynamics promoting
506 dioecy over hermaphroditism. *The American Naturalist*, 170, 485-491.
507 van Kleunen, M., Dawson, W., Bossdorf, O. & Fischer, M. (2014) The more the merrier:
508 Multi-species experiments in ecology. *Basic and Applied Ecology*, 15, 1-9.
509 van Kleunen, M., Dawson, W., Essl, F., Pergl, J., Winter, M., Weber, E., Kreft, H., Weigelt,
510 P., Kartesz, J., Nishino, M., Antonova, L.A., Barcelona, J.F., Cabezas, F.J., Cardenas,
511 D., Cardenas-Toro, J., Castano, N., Chacon, E., Chatelain, C., Ebel, A.L., Figueiredo,
512 E., Fuentes, N., Groom, Q.J., Henderson, L., Inderjit, Kupriyanov, A., Masciadri, S.,
513 Meerman, J., Morozova, O., Moser, D., Nickrent, D.L., Patzelt, A., Pelser, P.B.,
514 Baptiste, M.P., Poopath, M., Schulze, M., Seebens, H., Shu, W.-s., Thomas, J.,
515 Velayos, M., Wieringa, J.J. & Pysek, P. (2015) Global exchange and accumulation of
516 non-native plants. *Nature*, 525, 100-103.
517 Weigelt, P., König, C. & Kreft, H. (2017) The Global Inventory of Floras and Traits (GIFT)
518 database. Available at: http://gift.uni-goettingen.de

519 Whittaker, R.J. & Fernández-Palacios, J.M. (2007) *Island biogeography: ecology, evolution,*
520 *and conservation*. Oxford University Press, New York.
521 Zuur, A., Ieno, E., Walker, N., Saveliev, A. & Smith, G. (2009) *Mixed effects models and*
522 *extensions in ecology with R*. Springer, New York.

523 ## Data accessibility

524 The data supporting the results are available online from Dryad
525 doi:10.5061/dryad.89pd54f.

Table 1 Association of island and mainland occurrence of species with selfing ability based on the non-conservative dataset. Logistic regression coefficients of four multinomial logistic regressions testing how the likelihood of being island colonists and island endemics compared to mainland species depends on the self-compatibility or the autofertility index of species, scaled to a mean of zero and a standard deviation of one, their life history (annual/biennial or perennial), and the interaction between self-compatibility or autofertility and life history. Shown are the mean, the 2.5 % and the 97.5 % percentiles (95 % credible intervals) of a sample of 29,400 posterior distributions of each model parameter.

Multicategorical response variable with mainland species as a reference	Island colonist vs. mainland species	Island endemic vs. mainland species
Explanatory variables	Mean (2.5 %, 97.5 %)	Mean (2.5 %, 97.5 %)
Selfing-ability indices based on fruit set		
Analysis with self-compatibility (DIC = 1794.930)*		
Self-compatibility index	0.194 (0.042, 0.352)	0.282 (-0.100, 0.674)
Annual/biennial	1.058 (0.078, 2.093)	-2.126 (-6.017, 0.677)
Self-compatibility index x Annual/biennial	0.010 (-0.907, 0.883)	-2.565 (-5.838, 0.409)
Family (random effect)	0.804 (0.420, 1.407)	0.054 (0.024, 0.101)
Analysis with autofertility (DIC = 1297.600)†		
Autofertility index	0.359 (0.177, 0.542)	0.260 (-0.186, 0.672)
Annual/biennial	1.023 (0.200, 1.882)	-1.789 (-5.585, 0.930)
Autofertility index x Annual/biennial	-0.063 (-0.632, 0.519)	-2.596 (-6.616, 0.151)
Family (random effect)	0.533 (0.269, 0.959)	0.057 (0.024, 0.109)

Selfing-ability indices based on seed production		
Analysis with self-compatibility (DIC = 711.157)[‡]		
Self-compatibility index	0.300 (0.025, 0.582)	0.265 (-0.283, 0.845)
Annual/biennial	1.347 (0.180, 2.691)	1.101 (-1.195, 3.220)
Self-compatibility index x Annual/biennial	-0.691 (-1.921, 0.401)	-2.799 (-5.175, -0.642)
Family (random effect)	0.952 (0.33, 2.351)	0.090 (0.024, 0.221)
Analysis with autofertility (DIC = 487.485)[§]		
Autofertility index	0.373 (-0.019, 0.693)	0.227 (-0.507, 0.947)
Annual/biennial	0.931 (-0.146, 2.060)	0.213 (-2.361, 2.317)
Autofertility index x Annual/biennial	-0.309 (-1.160, 0.530)	-2.942 (-6.217, -0.787)
Family (random effect)	0.642 (0.211, 1.674)	0.095 (0.018, 0.276)

534 [*] $n_{\text{mainland species}} = 700$; $n_{\text{island colonists}} = 440$; $n_{\text{island endemics}} = 44$

535 [†] $n_{\text{mainland species}} = 517$; $n_{\text{island colonists}} = 315$; $n_{\text{island endemics}} = 36$

536 [‡] $n_{\text{mainland species}} = 207$; $n_{\text{island colonists}} = 199$; $n_{\text{island endemics}} = 13$

537 [§] $n_{\text{mainland species}} = 141$; $n_{\text{island colonists}} = 133$; $n_{\text{island endemics}} = 10$

538 **Table 2** Association of the occurrence of species on oceanic or continental islands with
539 selfing ability. Results of four logistic regressions testing how the occurrence of island
540 colonist species on oceanic islands depends on self-compatibility or the autofertility index of
541 species, scaled to a mean of zero and a standard deviation of one, their life history
542 (annual/biennial or perennial), and the interaction between self-compatibility or autofertility
543 and life history. Shown are the degrees of freedom (df), χ^2 values and *p* values from log-
544 likelihood ratio tests.

	Analysis with self-compatibility			Analysis with autofertility		
Selfing-ability indices based on fruit set						
Explanatory variables	df	χ^2	*p*	df	χ^2	*p*
Selfing ability index	1	0.849	0.357	1	0.002	0.964
Annual/biennial	1	6.884	0.009	1	4.792	0.029
Selfing ability x Annual/biennial	1	4.153	0.041	1	1.497	0.221
Random effect	sd	$sd_{residuals}$	*n*	sd	$sd_{residuals}$	*n*
Family	1.182	1.028	388	1.128	1.005	274
Selfing-ability indices based on seed production						
Explanatory variables	df	χ^2	*p*	df	χ^2	*p*
Selfing ability index	1	1.248	0.264	1	1.367	0.242
Annual/biennial	1	6.053	0.014	1	3.963	0.046
Selfing ability x Annual/biennial	1	3.833	0.050	1	2.014	0.156
Random effect	sd	$sd_{residuals}$	*n*	sd	$sd_{residuals}$	*n*

Family	0.725	1.055	171	1.138	0.953	113

545

546 **Table 3** Association of the occurrence of native vs. naturalized species on islands with selfing
547 ability. Results of four logistic regressions testing how the naturalization status of island
548 colonist species depends on the self-compatibility or the autofertility index of species, scaled
549 to a mean of zero and a standard deviation of one, their life history (annual/biennial or
550 perennial), and the interaction between self-compatibility or autofertility and life history.
551 Shown are the degrees of freedom (df), χ^2 values and *p* values from log-likelihood ratio tests.

	Analysis with self-compatibility			Analysis with autofertility		
Selfing-ability indices based on fruit set						
Explanatory variables	df	χ^2	*p*	df	χ^2	*p*
Selfing ability index	1	0.462	0.497	1	0.146	0.703
Annual/biennial	1	10.123	0.001	1	14.039	<0.001
Selfing ability x Annual/biennial	1	0.028	0.867	1	1.254	0.262
Random effect	sd	$sd_{residuals}$	*n*	sd	$sd_{residuals}$	*n*
Family	0.707	1.049	440	0.887	0.989	315
Selfing-ability indices based on seed production						
Explanatory variables	df	χ^2	*p*	df	χ^2	*p*
Selfing ability index	1	0.710	0.399	1	0.342	0.559
Annual/biennial	1	8.735	0.003	1	11.980	<0.001
Selfing ability x Annual/biennial	1	0.044	0.833	1	1.746	0.186
Random effect	sd	$sd_{residuals}$	*n*	sd	$sd_{residuals}$	*n*

Family	0.856	1.023	199	0.874	1.010	133

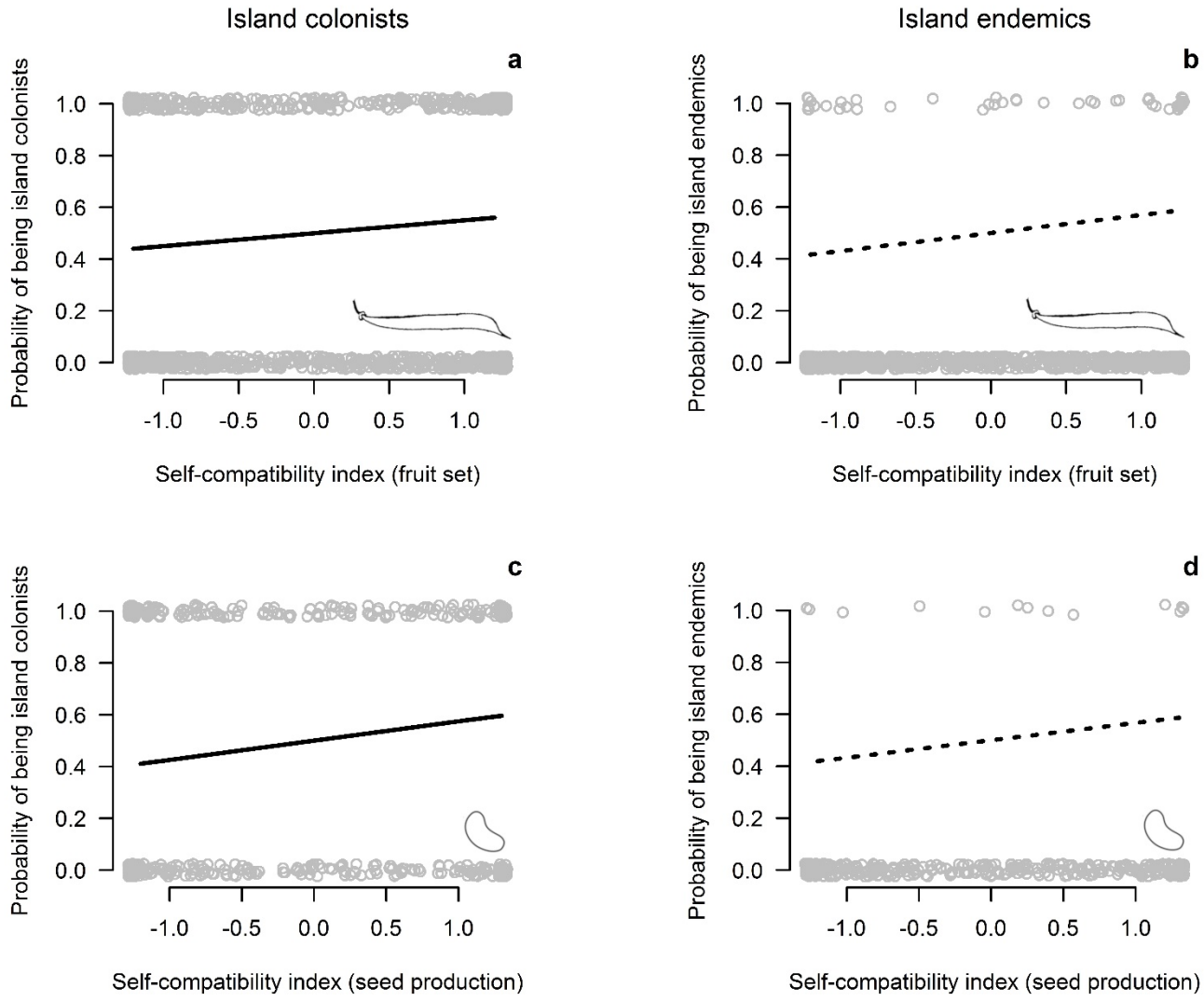
552

Figure legends

Figure 1 Island and mainland occurrence of plant species in relation to self-compatibility based on the non-conservative dataset. (a-d) Results of four multinomial logistic regressions testing how the occurrence of native and naturalized plant species on both the mainland and islands (island colonists) and only on islands (island endemics) compared to only on the mainland depends on self-compatibility of species. Each row corresponds to a model. Selfing ability of species was measured as (a-b) a self-compatibility index calculated based on fruit set (n=1184); (c-d) a self-compatibility index calculated based on seed production (n=419). Self-compatibility indices were standardized to a mean of zero and a standard deviation of one. Sample sizes refer to the total number of species from individual studies in the breeding system database. Solid curves and dotted curves indicate that the posterior distribution of the model parameter overlaps with zero and do not overlap with zero, respectively.

Figure 2 Island and mainland occurrence of plant species in relation to autofertility based on the non-conservative dataset. (a-d) Results of four multinomial logistic regressions testing how the occurrence of native and naturalized plant species on both the mainland and islands (island colonists) and only on islands (island endemics) compared to only on the mainland depends on autofertility of species. Each row corresponds to a model. Selfing ability of species was measured as (a-b) an autofertility index calculated based on fruit set (n=868); (c-d) an autofertility index calculated based on seed production (n=284). Autofertility indices were standardized to a mean of zero and a standard deviation of one. Sample sizes refer to the total number of species from individual studies in the breeding system database. Solid curves and dotted curves indicate that the posterior distribution of the model parameter overlaps with zero and do not overlap with zero, respectively.

Figure 3 Association of island and mainland occurrence of species with selfing ability based on the non-conservative dataset. (a-d) Posterior distribution of the logistic regression coefficients of four multinomial logistic regressions testing how the occurrence of native and naturalized plant species on both the mainland and islands (island colonists) and only on islands (island endemics) compared to only on the mainland (mainland species) depends on selfing ability of species. Selfing ability of species was measured (a) as an index of self-compatibility calculated based on fruit set (n=1184); (b) an index of autofertility calculated based on fruit set (n=868); (c) a self-compatibility index calculated based on seed production (n=419); (d) an autofertility index calculated based on seed production (n=284). Sample sizes, given in brackets, refer to the total number of species from individual studies in the breeding-system database.



587

588 **Figure 1** Island and mainland occurrence of plant species in relation to self-compatibility

589 based on the non-conservative dataset. (a-d) Results of four multinomial logistic regressions

590 testing how the occurrence of native and naturalized plant species on both the mainland and

591 islands (island colonists) and only on islands (island endemics) compared to only on the

592 mainland depends on self-compatibility of species. Each row corresponds to a model. Selfing

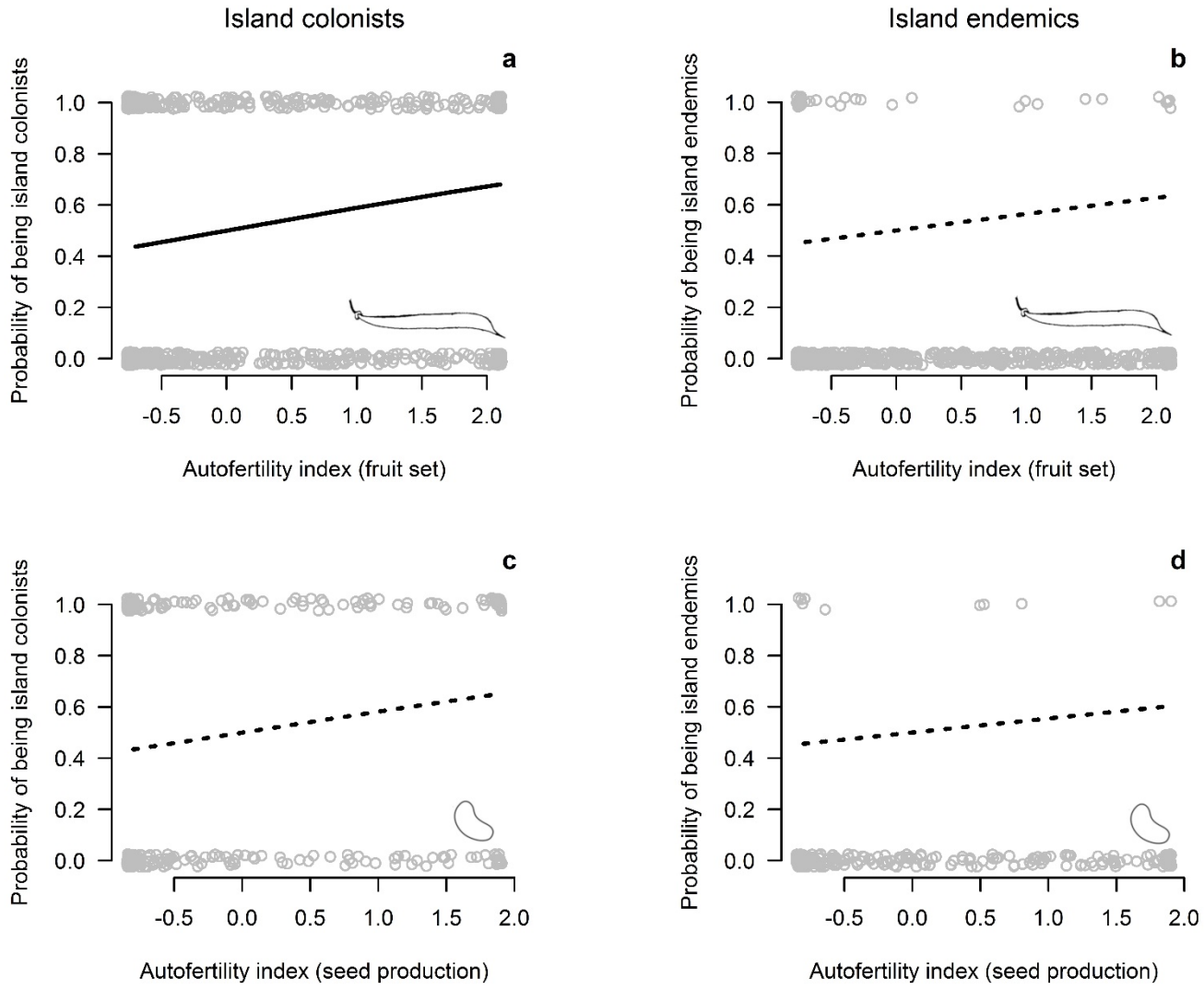
593 ability of species was measured as (a-b) a self-compatibility index calculated based on fruit

594 set (n=1184); (c-d) a self-compatibility index calculated based on seed production (n=419).

595 Self-compatibility indices were standardized to a mean of zero and a standard deviation of

596 one. Sample sizes refer to the total number of species from individual studies in the breeding

system database. Solid curves and dotted curves indicate that the posterior distribution of the model parameter overlaps with zero and do not overlap with zero, respectively.



599

600 **Figure 2** Island and mainland occurrence of plant species in relation to autofertility based on
601 the non-conservative dataset. (a-d) Results of four multinomial logistic regressions testing
602 how the occurrence of native and naturalized plant species on both the mainland and islands
603 (island colonists) and only on islands (island endemics) compared to only on the mainland
604 depends on autofertility of species. Each row corresponds to a model. Selfing ability of
605 species was measured as (a-b) an autofertility index calculated based on fruit set (n=868); (c-
606 d) an autofertility index calculated based on seed production (n=284). Autofertility indices
607 were standardized to a mean of zero and a standard deviation of one. Sample sizes refer to the
608 total number of species from individual studies in the breeding system database. Solid curves

609 and dotted curves indicate that the posterior distribution of the model parameter overlaps with
610 zero and do not overlap with zero, respectively.

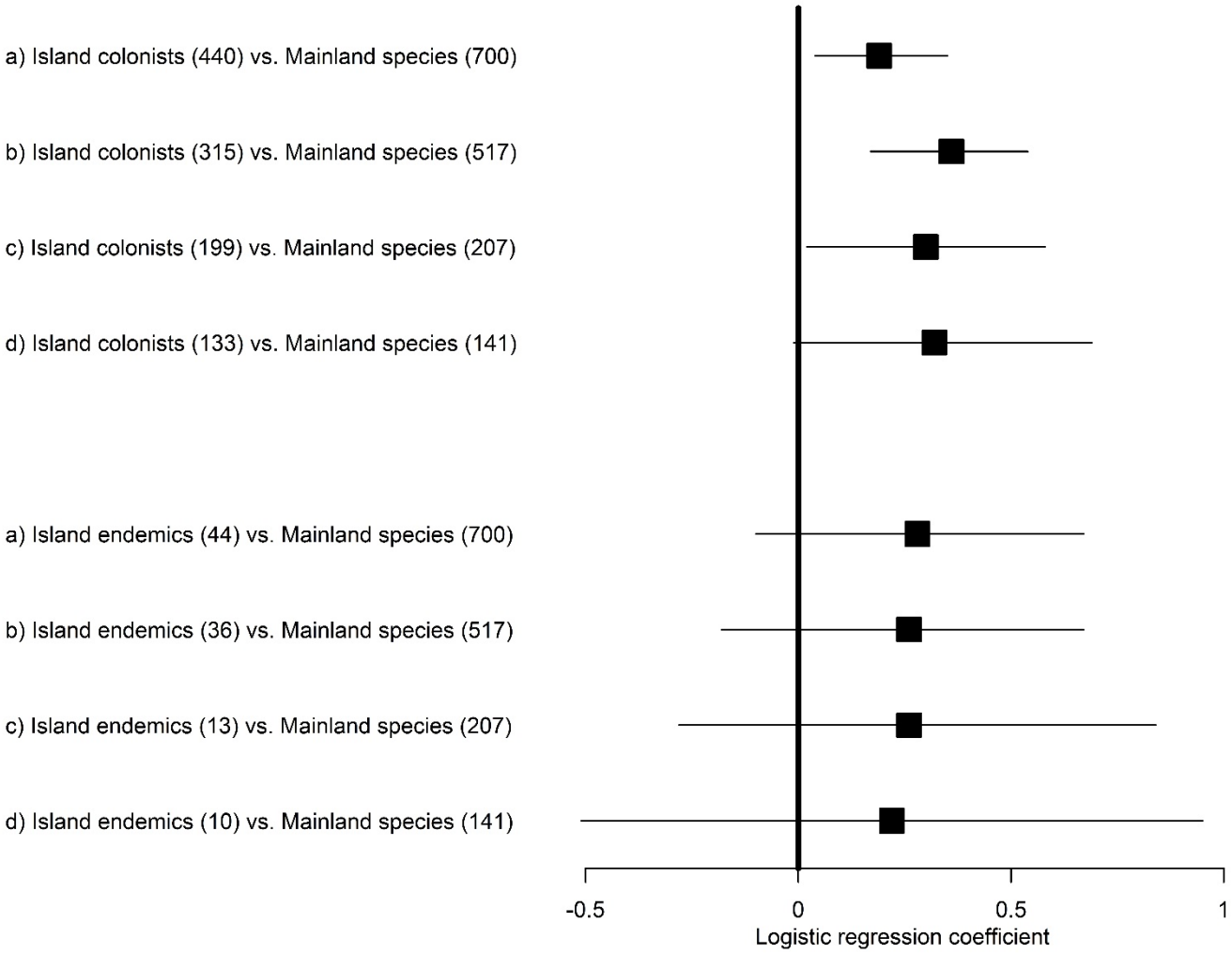


Figure 3 Association of island and mainland occurrence of species with selfing ability based on the non-conservative dataset. (a-d) Posterior distribution of the logistic regression coefficients of four multinomial logistic regressions testing how the occurrence of native and naturalized plant species on both the mainland and islands (island colonists) and only on islands (island endemics) compared to only on the mainland (mainland species) depends on selfing ability of species. Selfing ability of species was measured (a) as an index of self-compatibility calculated based on fruit set (n=1184); (b) an index of autofertility calculated based on fruit set (n=868); (c) a self-compatibility index calculated based on seed production (n=419); (d) an autofertility index calculated based on seed production (n=284). Sample sizes,

622 given in brackets, refer to the total number of species from individual studies in the breeding-
623 system database.